Записки научных
семинаров ПОМИ
Том 441, 2015 г.

R. Dakovic, M. Denker, $\boxed{\text{M. Gordin}}$

CIRCULAR UNITARY ENSEMBLES: PARAMETRIC MODELS AND THEIR ASYMPTOTIC MAXIMUM LIKELIHOOD ESTIMATES

Abstract. Parametrized families of distributions for the circular unitary ensemble in random matrix theory are considered which are connected to Toeplitz determinants and which have many applications in mathematics (for example to the longest increasing subsequences of random permutations) and physics (for example to nuclear physics and quantum gravity). We develop a theory for the unknown parameter estimated by an asymptotic maximum likelihood estimator, which, in the limit, behaves as the maximum likelihood estimator if the latter is well defined and the family is sufficiently smooth. They are asymptotically unbiased and normally distributed, where the norming constants are unconventional because of long range dependence.

§1. Introduction

The purpose of this note is to initiate a statistical analysis of parametrized exponential families of distributions on the group of unitary matrices which are dominated by the Haar measure. Statistical analysis of the circular unitary ensemble (CUE) is not entirely new in general, but the theory of maximum likelihood estimates is and has been first investigated by one of the authors in [29].

In order to motivate the approach here we recall basic facts from random matrix theory and explain which role the parametrized families play within the theory. Random matrix theory (RMT) is used in different branches of science since it was introduced by Wishart in the late 1920s. It has been successfully applied to an extraordinarily large variety of problems in fields

Key words and phrases: circular unitary ensemble, Toeplitz determinant, maximum likelihood estimator, normal distribution, long range dependence.

Rada Dakovic acknowledges the support of the Center for Statistics, University of Göttingen. Manfred Denker was supported by nùmero 158/2012 de PesquisadorVisitante Especial de CAPES. Mikhail Gordin supported by the RFBR grant 13-01-00256 and by grant NSh: 2504.2014.1.

as diverse as multivariate statistics [2, 32], and more recently [6, 24], harmonic analysis on groups [10], combinatorics [3], nuclear physics [31], quantum gravity [15], wireless communications [41], to name a few important applications.

Although random matrices were first encountered by Weyl [42, 43] in connection with the integration over the unitary group, the explicit study of their properties began in 1928 with Wishart [48], who obtained the joint distribution of sample variances and covariances from multivariate normal populations. After a relatively slow start, the investigation of random matrix ensembles intensified in the 1950s, when Wigner [45, 46] proposed to use RMT to characterize certain properties of complex many-body systems such as heavy nuclei, complex atoms and molecules. Despite the fact that the Gaussian ensembles, introduced in Wigner [46, 47], allow a wide range of applications, they have the unpleasant property of being defined on the non-compact space of matrices. Consequently there is no way assigning the same weight to every matrix and hence matrices representing different quantum systems cannot be treated in the same manner. To avoid this deficiency, Dyson [17] modified Wigner's treatment of a nucleus and defined three ensembles, similar to Gaussian, but mathematically simpler to deal with. The three circular ensembles (orthogonal, unitary and symplectic) are defined as subsets of the set $U(n)$ of all $n \times n$ unitary matrices. They are applied in a broad range of models (see [18, 31] for details). Nevertheless, there exist physical systems that could be described by random unitary matrices which do not share the statistical properties of Dyson's circular ensembles. As mentioned in Muttalib, Ismail [33], numerical studies related to disordered conductors [23] and periodically driven systems [28], exhibit statistical behavior different from that of Dyson's models. Therefore, generalizations of circular ensembles of Dyson were introduced in [33, 34].

We shall develop a rigorous statistical treatment of generalized circular unitary ensembles (GCUE), where the attention is restricted to the unitary case since the analytical tools are developed best in this case; in particular, the asymptotic theory of Toeplitz determinants can be used. The investigation is motivated by numerous examples from mathematics and physics. Examples from physics include chaotic scattering (Jalabert, Pichard [23]), conductance in mesoscopic systems (Beenakker [7]) and periodically driven systems (Haake [20]). Related models from mathematics arise in the theory of orthogonal polynomials on the unit circle (Ismail, Witte [22], Simon [37]), the theory of Toeplitz determinants (Adler, van

Moerbeke [1], Borodin, Okounkov [9]), and studies of the length of the longest increasing subsequence in a random permutation (Johansson [26], Baik, Deift, Johansson [3], etc). The attention of physicists and mathematicians is often centered around the asymptotic behavior of the spectra of random matrices rather than the structure of random matrices itself. Therefore, our generalization of Dyson's circular unitary ensemble will be defined in terms of the density of the joint eigenvalue distribution. Although the estimation problem can be dealt with within a general non-parametric framework, we restrict our attention to the case of densities belonging to exponential families of probability distributions, since numerous examples discussed in the physics literature allow this interpretation (see the end of Section 2). Additionally, remarkable analytical tools from the theory of exponential families can be applied which are not available in the non-parametric setup.

This estimation problem has been first studied by the first author in her PhD dissertation [29]. It contains the central limit theorems for sufficient statistics of the four classical β-ensembles and the generalized circular unitary ensemble. Central limit theorems were obtained (in a broader context) for β-Hermite and Dyson's circular ensembles in Johansson [25, 27], Dumitriu, Edelman [16], Diaconis, Evans [14]. Corresponding theorems for sufficient statistics of β-Laguerre, β-Jacobi, the Cauchy unitary ensemble are obtained in [29] which also forms the basis of the result in this note. All results on the asymptotic normality of the sufficient statistics, as well as the results on the behavior of the asymptotic maximum likelihood estimator defined in Sections 4 and 5 seem to be new. Since this paper had been written in the years 2007 to 2009 (shortly after the PhD thesis [29] had been defended) several papers appeared which are in some way related to the present one. We mention [5] where correlation functions of parametrized families are considered, [36] where similar techniques as here are used to derive asymptotic distributions. The book of Bai and Silverstein [6] shows a wide range of applications of large dimensional random matrices and gives a basic account of theoretical results. The present work adds to this endeavor.

This paper seems to be one of the first attempts to apply parametric statistical methods to analyze spectra of random matrices in the non-Gaussian case. Classical estimation theory of i.i.d. samples is not applicable in this setting since our analysis involves inference for samples of *n strongly dependent* observations. We believe that the results in this note can serve as another starting point for a rich research area in high-dimensional asymptotic statistics.

A short description of the models considered and a summary of our results follows.

Let U be a unitary matrix taken at random with respect to the normalized Haar measure from the unitary group $U(n)$, $n \geq 1$. The random eigenvalues of U are denoted by $(e^{iZ_1}, \dots, e^{iZ_n})$, where $Z_k \in [0, 2\pi)$ for $k = 1, \dots, n$. The joint distribution of eigenphases (called the Weyl measure) is absolutely continuous with respect to the Lebesgue measure on $[0, 2\pi)^n$ and its density is

$$p_n(\boldsymbol{\zeta}) = C_n^{-1} \prod_{1 \leqslant k < l \leqslant n} |e^{i\zeta_k} - e^{i\zeta_l}|^2, \qquad \boldsymbol{\zeta} \in [0, 2\pi)^n, \tag{1}$$

where C_n is the normalizing constant (see Weyl [42, 43]).

In this paper we introduce a generalization of (1) letting the probability density of n eigenphases be of the form

$$p_{\boldsymbol{\theta};\, n}(\boldsymbol{\zeta}) = C_n^{-1}(\boldsymbol{\theta}) \prod_{k=1}^{n} w_{\boldsymbol{\theta}}(e^{i\zeta_k}) \prod_{1 \leqslant k < l \leqslant n} |e^{i\zeta_k} - e^{i\zeta_l}|^2, \qquad \boldsymbol{\zeta} \in [0, 2\pi)^n,$$

where $\boldsymbol{\theta} = (\theta_1, \dots, \theta_r) \in \Theta \subset \mathbb{R}^r$ is an r-dimensional unknown, and hence an estimable parameter. $C_n(\boldsymbol{\theta})$ is the normalizing constant, and $w_{\boldsymbol{\theta}}$ is a suitable weight function on the unit circle $\mathbb{T}$. The theory of Toeplitz determinants and complex analysis of several complex variables turns out to be helpul deriving properties of the asymptotic maximum likelihood estimator of $\boldsymbol{\theta}$, and showing that the Fisher information contained in the sample $\boldsymbol{\zeta} = (\zeta_1, \dots, \zeta_n)$ stays bounded as the dimension parameter n increases provided the weight function $w_{\boldsymbol{\theta}}$ does not involve n. To overcome this deficiency, we consider the third-order phase transition model with varying weight introduced in Gross, Witten [19], where the probability density of n eigenphases is equal to

$$p_{\gamma;\, n}(\boldsymbol{\zeta}) = C_n^{-1}(\gamma) \prod_{k=1}^{n} \exp\left\{\gamma n \cos \zeta_k\right\} \prod_{1 \leqslant k < l \leqslant n} |e^{i\zeta_k} - e^{i\zeta_l}|^2, \qquad \boldsymbol{\zeta} \in [0, 2\pi)^n \tag{2}$$

with normalizing constant $C_n(\gamma)$ and parameter $\gamma > 0$. We show that the true parameter value γ_0 can be estimated consistently by an asymptotic maximum likelihood estimator, provided that $\gamma_0 \in (0, 1 - \epsilon)$, $\epsilon > 0$. Additionally, we obtain that the asymptotic maximum likelihood estimator $\widehat{\gamma}_n$

of γ_0 is asymptotically normal in the sense that

$$n(\widehat{\gamma}_n - \gamma_0) \xrightarrow{\mathscr{D}} \mathscr{N}(0,2), \qquad n \to \infty,$$

if $\gamma_0 \in (0, 1-\epsilon)$, $\epsilon > 0$. This result is in sharp contrast to classical estimation theory of i.i.d. samples under regularity conditions where the variance of the maximum likelihood estimator is of order $O(1/n)$.

The outline of this paper is as follows. In Section 2, we state basic facts about Dyson's circular unitary ensemble and formally introduce the generalized circular unitary ensemble (GCUE) extending the model CUE. In Section 3 the asymptotic normality of the sufficient statistics for GCUE is established. The asymptotic maximum likelihood procedure is applied to estimate the parameters of the joint eigenvalue distribution of GCUE. The properties of the estimators are derived in Section 4. In Section 5 we analyze the model with varying weight from Gross, Witten [19]. Section 6 contains a discussion of open questions.

§2. Circular unitary ensemble and generalizations

The *circular unitary ensemble* (CUE) is the group of unitary matrices endowed with the normalized Haar measure. As mentioned in the introduction, the joint distribution of n eigenvalues of a matrix taken from CUE is absolutely continuous with respect to the Lebesgue measure on $\mathbb{T}^n$ and the density function of phases is given by (1).

The density can be rewritten using the Vandermonde determinant

$$\Delta(\boldsymbol{\zeta}) = (-1)^{\frac{n(n-1)}{2}} \prod_{1 \leqslant k < l \leqslant n} (e^{i\zeta_k} - e^{i\zeta_l})$$
$$= \begin{vmatrix} 1 & 1 & \dots & 1 \\ e^{i\zeta_1} & e^{i\zeta_2} & \dots & e^{i\zeta_n} \\ \dots & \dots & \dots & \dots \\ e^{i(n-1)\zeta_1} & e^{i(n-1)\zeta_2} & \dots & e^{i(n-1)\zeta_n} \end{vmatrix}$$

and this shows that the eigenphases $(Z_1, \dots, Z_n)$ with joint probability density (1) exhibit strong repulsive dependence. This repulsiveness property allows applications of CUE in a variety of models in physics and mathematics. However, numerical studies of periodic quantum systems and disordered conductors, scattering of plane waves within an irregularly shaped domain (see [18,33]) have a statistical behavior different from that of CUE. A new analytical model, containing CUE is thus considered.

Definition 1. *The* generalized circular unitary ensemble *is the ensemble of random unitary matrices whose joint distribution of n eigenphases is given by the density*

$$p_{\boldsymbol{\theta};\,n}(\boldsymbol{\zeta}) = C_n^{-1}(\boldsymbol{\theta}) \prod_{k=1}^{n} w_{\boldsymbol{\theta}}(e^{i\zeta_k}) \prod_{1 \leqslant k < l \leqslant n} |e^{i\zeta_k} - e^{i\zeta_l}|^2, \qquad \boldsymbol{\zeta} \in [0, 2\pi)^n, \quad (3)$$

where $\boldsymbol{\theta} = (\theta_1, \dots, \theta_r)$ *is an* r*-dimensional parameter,* $C_n(\boldsymbol{\theta})$ *is the normalizing constant and*

$$w_{\boldsymbol{\theta}}(e^{i\zeta}) = \exp\Big\{\sum_{j=1}^{r} \theta_j V_j(e^{i\zeta})\Big\}, \quad \zeta \in [0, 2\pi], \qquad (4)$$

with real-valued functions V_j *having Fourier coefficients* $\{\widehat{V_j}(k)\}_{k \in \mathbb{Z}}$ *satisfying the condition*

$$E_j = \sum_{k \in \mathbb{N}} k \left| \widehat{V_j}(k) \right|^2 < \infty, \qquad 1 \leqslant j \leqslant r. \qquad (5)$$

We note that the function $w_{\boldsymbol{\theta}}$ is integrable on the unit circle if the condition (5) is satisfied. This fact is obtained from the first proposition (for a proof see [37]).

Proposition 2.1. *Let* $V \in L^1(\mathbb{T})$ *be a real-valued function on the unit circle with Fourier coefficients* $\{\widehat{V}(k)\}_{k \in \mathbb{Z}}$*, such that*

$$E = \sum_{k \in \mathbb{N}} k \left| \widehat{V}(k) \right|^2 < \infty.$$

Then the function $e^{i\zeta} \mapsto w(e^{i\zeta}) = \exp V(e^{i\zeta})$*,* $\zeta \in [0, 2\pi)$*, belongs to* $L^1(\mathbb{T})$*.*

It follows from this proposition that $\exp(\theta_j V_j) \in L^1(\mathbb{T})$ for every $\theta_j \in \mathbb{R}$, $1 \leqslant j \leqslant r$, and consequently $w_{\boldsymbol{\theta}} \in L^1(\mathbb{T})$ for every $\boldsymbol{\theta} \in \mathbb{R}^r$ whenever the condition (5) is satisfied. Since $|\Delta(\boldsymbol{\zeta})|^2$ is bounded on $\mathbb{T}^n$, the density (3) is well defined.

If the condition (5) is satisfied, the strong Szegő theorem for Toeplitz determinants

$$D_n[f] = \det[\widehat{f}(k-l)]_{0 \le k,l < n} = \begin{vmatrix} \widehat{f}(0) & \widehat{f}(-1) & \dots & \widehat{f}(-n+1) \\ \widehat{f}(1) & \widehat{f}(0) & \dots & \widehat{f}(-n+2) \\ \dots & \dots & \dots & \dots \\ \widehat{f}(n-1) & \widehat{f}(n-2) & \dots & \widehat{f}(0) \end{vmatrix}$$

($f \in L^1(\mathbb{T})$) can be used to derive statistical properties of the asymptotic MLE for the parameter $\boldsymbol{\theta}$ in (3). The key connection of the GCUE with the theory of Toeplitz determinants is based on the following determinantal identity due to Heine and Szegő; in particular, the explicit expression for the normalizing constant $C_n(\boldsymbol{\theta})$ can be deduced (see [27], equation (1.4)).

Lemma 2.2 (Szegő [39]). *Let $f \in L^1(\mathbb{T})$ be a function on the unit circle with Fourier coefficients $\{\widehat{f}(k)\}_{k\in\mathbb{Z}}$. Then the following identity holds*

$$\frac{1}{(2\pi)^n n!} \int\limits_{[0,2\pi)^n} \prod_{k=1}^{n} f(e^{i\zeta_k})\, |\Delta(\boldsymbol{\zeta})|^2 \, d\boldsymbol{\zeta} = D_n[f]. \tag{6}$$

From this we can see that $C_n(\boldsymbol{\theta})$ in (3) is a multiple of the Toeplitz determinant with respect to the symbol $w_{\boldsymbol{\theta}}$,

$$C_n(\boldsymbol{\theta}) = (2\pi)^n n!\, D_n[w_{\boldsymbol{\theta}}]. \tag{7}$$

Below we also use the notation

$$C_n(\boldsymbol{\theta} + i\boldsymbol{t}) = (2\pi)^n n!\, D_n[f_{\boldsymbol{\theta},\boldsymbol{t}}] \qquad \boldsymbol{\theta}, \boldsymbol{t} \in \mathbb{R}^r,$$

when

$$f_{\boldsymbol{\theta},\boldsymbol{t}}(\boldsymbol{\zeta}) = \prod_{k=1}^{n} \left[\sum_{j=1}^{r} (\theta_j + it_j) V_j(e^{i\zeta_k}) \right]$$

($\boldsymbol{t} = (t_j)_{1\le j\le r}$, $\boldsymbol{\theta} = (\theta_j)_{1\le j\le r}$, $\boldsymbol{\zeta} = (\zeta_k)_{1\le k\le n}$). The smoothness condition on f' was relaxed to the condition that $\log f \in L_1(\mathbb{T})$ in [21] and the final form we shall need can be found in [25, Theorem 2.2] (attributed to Onsager and Szegő).

Theorem 2.3. *Let $g \in L^1(\mathbb{T})$ be a complex-valued function on $\mathbb{T}$ with Fourier coefficients $\{\widehat{g}(k)\}_{k\in\mathbb{Z}}$. Assume that*

$$\sum_{k\in\mathbb{Z}} |k|\, |\widehat{g}(k)|^2 < \infty.$$

Then

$$D_n[\exp g] = \exp\left\{ n\widehat{g}(0) + \sum_{k\in\mathbb{N}} k\, \widehat{g}(k)\, \widehat{g}(-k) + o(1) \right\} \tag{8}$$

as $n \to \infty$.

The identity (8) is used to establish the CLT for the sufficient vector statistic of GCUE and to derive the asymptotic MLE of the parameter $\boldsymbol{\theta}$. The approach by (6) and Theorem 2.3 has been used by Johansson [26] to obtain the asymptotic normality of a broad class of linear statistics for CUE.

We conclude this section with a motivation for the models described in (3). Although a non-parametric setup can be considered, the class of models here covers most practical objectives. Moreover, statistical analysis for these models can widely be developed in the framework of the theory of exponential families: the models allow a reduction of data by sufficiency, and the normalizing constant $C_n(\boldsymbol{\theta})$ is an analytical function of its parameter $\boldsymbol{\theta} \in \mathbb{R}^r$ for arbitrary fixed $n \in \mathbb{N}$. In addition to these properties, we mention that the parameters $\theta_1, \ldots, \theta_r$ may have a clear physical interpretation, and their estimation often is of high importance for understanding the system's behavior. As mentioned in Muttalib, Ismail [33], the introduction of the weight function $w_{\boldsymbol{\theta}}$ may come from any system-dependent physical constraint, since the measure under consideration may depend on various physical parameters.

The main examples of the weight function $w_{\boldsymbol{\theta}}$ in (4), include

$$\begin{aligned}
w_\theta^{(1)}(e^{i\zeta}) &= \exp\left(\theta \cos\zeta\right), && \zeta \in [0, 2\pi], \\
w_{\boldsymbol{\theta}}^{(2)}(e^{i\zeta}) &= \exp\left(\theta_1 \cos\zeta + \theta_2 \cos 2\zeta\right), && \zeta \in [0, 2\pi], \\
w_\theta^{(3)}(e^{i\zeta}) &= (1 + \rho^2 - 2\rho\cos\zeta)^\theta, && \zeta \in [0, 2\pi].
\end{aligned}$$

The weight function $w_\theta^{(1)}$ and the corresponding system of orthogonal polynomials arose from the studies on the length of the longest increasing subsequence of a random permutation in Baik, Deift, Johansson [3, p. 1123], and random matrix models in Gross, Witten [19], Periwal, Shewitz [35]. Recursion relations for Toeplitz determinants of this symbol can be found in Borodin [11], or Adler, van Moerbeke [1, Section 2], while the properties of orthogonal polynomials with respect to this weight (called the Bessel weight) appear in detail in Ismail, Witte [22]. It should be noted here that in the literature on circular statistics, $w_\theta^{(1)}(e^{i\zeta})$ is known as the density of the von Mises–Fisher distribution on the 2-dimensional sphere (see e.g. Jammalamadaka, SenGupta [24]). The weight $w_{\boldsymbol{\theta}}^{(2)}$ arose in the studies of the longest increasing subsequence in a random odd permutation (see Tracy, Widom [40]). Rational recursion relations for the respective Toeplitz determinants are described in detail in Adler, van Moerbeke [1].

For the discussion of the weight $w_\theta^{(3)}$ and their Toeplitz determinants the reader may consult Borodin [11], Borodin, Okounkov [9], and Adler, van Moerbeke [1].

Remark 2.4. An important application of GCUE is the theory of log-potential gases. In this framework the density (3) coincides with the Boltzmann factor of a one-dimensional Coulomb gas consisting of n particles free to move on the unit circle in an external field with potential $-\sum_{j=1}^{r}\theta_j V_j(\zeta)/2$ at temperature $1/2$. Further details regarding this connection are contained in Forrester [18].

Remark 2.5. We note here that due to Weyl's integration formula an equivalent definition of GCUE can be given by considering the probability measures with densities

$$\exp\Big\{\sum_{j=1}^{r}\theta_j \operatorname{tr}(V_j(U))\Big\}$$

on the unitary group $U(n)$, where the functions V_j are as given in Definition 1, considered as a class function on $U(n)$ (see [13, p. 92ff]). More generally and in the same spirit, measures with densities

$$\exp\{\operatorname{tr}(V(U))\}$$

can be considered where $V : \mathbb{T} \to \mathbb{R}$ is such that $\exp(V(\cdot))$ is integrable on $\mathbb{T}$. The estimation of the function V can be studied in the nonparametric setup. Ensembles of this form can be considered dominant in RMT and the theory of log-potential systems (see e.g. Deift [13]).

§3. Asymptotic distribution of the sufficient statistic

The vector statistic

$$\boldsymbol{V}^{(n)}(\boldsymbol{Z}) = \Big(\sum_{k=1}^{n} V_1(e^{iZ_k}), \dots, \sum_{k=1}^{n} V_r(e^{iZ_k})\Big) \tag{9}$$

in Definition 1 is clearly sufficient, where $\boldsymbol{Z} = (Z_1, \dots, Z_n)$ (we suppress the dependence of Z_l on n since no confusion can occur) is a sample of exchangeable random variables with probability density function

$$p_{\boldsymbol{\theta};\,n}(\boldsymbol{\zeta}) = C_n^{-1}(\boldsymbol{\theta}) \prod_{k=1}^{n} w_{\boldsymbol{\theta}}(e^{i\zeta_k}) \prod_{1\leqslant k<l\leqslant n} |e^{i\zeta_k} - e^{i\zeta_l}|^2, \qquad \boldsymbol{\zeta} \in [0, 2\pi)^n.$$

We shall prove a central limit theorem (CLT) for this statistics as $n \to \infty$. This result is of special importance when comparing models with experiments (physicists are mainly interested in those features of the statistical model that tend to definite limits as $n \to \infty$). The method of proof is based on the strong Szegő theorem and the identity (7), which will be used to prove pointwise convergence of the characteristic functions to the corresponding limit characteristic function of a normal distribution.

In the case of Dyson's circular unitary ensemble

$$p_{\mathbf{0};\,n}(\boldsymbol{\zeta}) = C_n^{-1}(\mathbf{0}) \prod_{1 \leqslant k < l \leqslant n} |e^{i\zeta_k} - e^{i\zeta_l}|^2, \qquad \boldsymbol{\zeta} \in [0, 2\pi]^n,$$

the central limit theorem was obtained for a statistics of the form

$$\sum_{k=1}^{n} f(e^{iZ_k}),$$

where $f \in L^1(\mathbb{T})$ is assumed to be real-valued function such that

$$\sum_{k \in \mathbb{N}} k|\widehat{f}(k)|^2 < \infty. \tag{10}$$

This result together with a superexponential rate of convergence for f is due to Johansson [25, Theorem 2.6]. For further developments of this subject see Soshnikov [38] and Diaconis, Evans [14].

In order to state the CLT for GCUE, we need to define a non-negative bilinear form on the space $H_2^{1/2}$ of real-valued functions $f \in L^1(\mathbb{T})$ satisfying the condition (10). The form is defined by

$$\langle f, g \rangle_{1/2} = \sum_{k \in \mathbb{Z}} |k|\, \widehat{f}(k)\widehat{g}(-k), \quad f, g \in H_2^{1/2}, \tag{11}$$

and turns $H_2^{1/2}$ into a Besov potential space. For details we refer the reader to [14] and references therein. Once the form (11) is defined, we can state the theorem

Theorem 3.1. *Consider the sequence of probability density functions* $\{p_{\boldsymbol{\theta};\,n}(\boldsymbol{\zeta})\}_{n \geqslant 2}$ *in* (3), *where the real-valued functions* V_j *belong to* $H_2^{1/2}$ *for each* $1 \leqslant j \leqslant r$. *Assume in addition that*

$$\widehat{V_j}(0) = 0, \qquad 1 \leqslant j \leqslant r.$$

Then, for fixed $\boldsymbol{\theta} \in \mathbb{R}^r$*, the vector statistic* $\boldsymbol{V}^{(n)}(\boldsymbol{Z})$ *in* (9) *has asymptotically a normal distribution with mean*

$$\boldsymbol{\mu} = \Big(\sum_{j=1}^{r} \theta_j \langle V_1, V_j \rangle_{1/2}, \ldots, \sum_{j=1}^{r} \theta_j \langle V_r, V_j \rangle_{1/2} \Big), \tag{12}$$

and covariance matrix

$$\Sigma = \begin{pmatrix} \langle V_1, V_1 \rangle_{1/2} & \langle V_1, V_2 \rangle_{1/2} & \ldots & \langle V_1, V_r \rangle_{1/2} \\ \langle V_2, V_1 \rangle_{1/2} & \langle V_2, V_2 \rangle_{1/2} & \ldots & \langle V_2, V_r \rangle_{1/2} \\ \ldots & \ldots & \ldots & \ldots \\ \langle V_r, V_1 \rangle_{1/2} & \langle V_r, V_2 \rangle_{1/2} & \ldots & \langle V_r, V_r \rangle_{1/2} \end{pmatrix}. \tag{13}$$

Proof. The asymptotic distribution of the statistic $\boldsymbol{V}^{(n)}(\boldsymbol{Z})$ is obtained from the convergence of its characteristic function

$$\phi_n(\boldsymbol{t}) = \mathrm{E}_n \exp \Big\{ \sum_{j=1}^{r} \sum_{k=1}^{n} i t_j V_j(e^{iZ_k}) \Big\}, \qquad \boldsymbol{t} = (t_j)_{1 \le j \le r} \in \mathbb{R}^r,$$

where E_n denotes the expectation with respect to the probability (3). Substituting the density $p_{\boldsymbol{\theta};\,n}(\cdot)$ into the mathematical expectation yields

$$\begin{aligned} \phi_n(\boldsymbol{t}) &= C_n^{-1}(\boldsymbol{\theta}) \int_{[0,2\pi]^n} |\Delta(\boldsymbol{\zeta})|^2 \prod_{k=1}^{n} \exp \Big\{ \sum_{j=1}^{r} (\theta_j + i t_j)\, V_j(e^{i\zeta_k}) \Big\} d\boldsymbol{\zeta} \\ &= \frac{C_n(\boldsymbol{\theta} + i\boldsymbol{t})}{C_n(\boldsymbol{\theta})} \end{aligned}$$

The identity (7) and Ibragimov's version of the strong Szegő theorem (Theorem 2.3) imply that the above expression has the limit

$$\begin{aligned} \lim_{n\to\infty} \phi_n(\boldsymbol{t}) &= \exp \Big\{ -\sum_{k \in \mathbb{N}} k \Big(\sum_{j=1}^{r} \theta_j\, \widehat{V}_j(k) \Big) \Big(\sum_{j=1}^{r} \theta_j\, \widehat{V}_j(-k) \Big) \Big\} \\ &\quad \times \exp \Big\{ \sum_{k \in \mathbb{N}} k \Big(\sum_{j=1}^{r} (\theta_j + i t_j)\, \widehat{V}_j(k) \Big) \Big(\sum_{j=1}^{r} (\theta_j + i t_j)\, \widehat{V}_j(-k) \Big) \Big\}. \end{aligned}$$

Rewriting the expression on the right-hand side, we observe that

$$\lim_{n\to\infty} \phi_n(\boldsymbol{t}) = \exp\Big[\sum_{j=1}^{r} it_j\Big(\sum_{l=1}^{r}\theta_l \,\langle V_j, V_l\rangle_{1/2}\Big)\Big] \\ \times \exp\Big[-\frac{1}{2}\sum_{j,l=1}^{r} t_j t_l \,\langle V_j, V_l\rangle_{1/2}\Big],$$

where the limiting function coincides with the characteristic function of the multivariate normal distribution with mean and covariance matrix as specified in (12) and (13). This completes the proof of the theorem. □

Remark 3.2. Notice that as in the case of CUE (see [25, Theorem 2.6], [14, Theorem 4.1], or [38, Theorem 1], also the result in [6, Theorem 1.1]). Theorem 3.1 proves the convergence

$$\boldsymbol{V}^{(n)}(\boldsymbol{Z}) \xrightarrow{\mathscr{D}} \mathscr{N}(\boldsymbol{\mu}, \Sigma), \quad n \to \infty,$$

without normalization by $\sqrt{n}$. Moreover, the components of the vector statistics $\boldsymbol{V}^{(n)}(\boldsymbol{Z})$ are asymptotically independent if and only if the functions $V_j, 1 \leqslant j \leqslant r$, are orthogonal in the space $H_2^{1/2}$, i.e. $\langle V_j, V_l\rangle_{1/2} = 0$ for $j \neq l$. In that case, the covariance matrix (13) obtains the diagonal structure

$$\Sigma = \mathrm{diag}\Big\{\langle V_1, V_1\rangle_{1/2}, \ldots, \langle V_r, V_r\rangle_{1/2}\Big\}.$$

§4. Asymptotic Maximum Likelihood Estimation

The aim of this section is to construct the asymptotic maximum likelihood estimator $\widehat{\boldsymbol{\theta}}^{(n)}$ for the parameter $\boldsymbol{\theta}$ of the generalized circular unitary ensemble. The maximum likelihood estimator $\tilde{\boldsymbol{\theta}}^{(n)}$ is defined as the solution of the system of equations

$$\frac{\partial}{\partial\theta_j}\log C_n(\tilde{\boldsymbol{\theta}}^{(n)}) = \sum_{k=1}^{n} V_j(e^{i\zeta_k}), \quad 1 \leqslant j \leqslant r, \tag{14}$$

and is not available in a closed form. We show that

$$\frac{\partial}{\partial\theta_j}\log C_n(\boldsymbol{\theta}) \xrightarrow[n\to\infty]{} \sum_{l=1}^{r}\theta_l \,\langle V_j, V_l\rangle_{1/2}, \quad 1 \leqslant j \leqslant r,$$

uniformly over $\boldsymbol{\theta} \in K$, $K \subset \mathbb{R}^r$, K compact, and replace (14) by the system of *linear* equations

$$\sum_{l=1}^{r} \widehat{\theta}_l^{(n)} \langle V_j, V_l \rangle_{1/2} = \sum_{k=1}^{n} V_j(e^{i\zeta_k}), \quad 1 \leqslant j \leqslant r,$$

which defines the asymptotic MLE $\widehat{\boldsymbol{\theta}}^{(n)}$ in a closed form. The estimator $\widehat{\boldsymbol{\theta}}^{(n)}$ is asymptotically unbiased, but not consistent. Below we indicate why a consistent estimation of the parameter $\boldsymbol{\theta}$ is not possible, unless the dimension parameter n is included into the weight function $w_{\boldsymbol{\theta}}$.

Since $p_{\boldsymbol{\theta},n}(\boldsymbol{\zeta})$ is the density of the distribution from an exponential family for each $n \geqslant 2$, it follows from [12, Corollary 2.6], that the derivatives of the function $\boldsymbol{\theta} \mapsto \lambda_n(\boldsymbol{\theta}) := C_n(\boldsymbol{\theta})$ may be obtained by differentiation under the integral sign. Additionally, $\lambda_n(\boldsymbol{\theta})$ admits an analytical continuation to $\mathbb{C}^r$ by the identity

$$\tilde{\lambda}_n(\boldsymbol{\theta}) = \int_{[0,2\pi)^n} |\Delta(\boldsymbol{\zeta})|^2 \prod_{k=1}^{n} w_{\boldsymbol{\theta}}(e^{i\zeta_k}) d\boldsymbol{\zeta}, \quad \boldsymbol{\theta} \in \mathbb{C}^r.$$

Assume that the conditions of Theorem 3.1 are satisfied and consider the sequence of Toeplitz determinants $\{D_n[w_{\boldsymbol{\theta}}]\}_{n\geqslant 2}$ with respect to the symbol $w_{\boldsymbol{\theta}}$, $\boldsymbol{\theta} \in \mathbb{C}^r$. Since the functions $V_j, 1 \leqslant j \leqslant r$ are fixed, the Toeplitz determinants $D_n[w_{\boldsymbol{\theta}}]$ become functions of the parameter $\boldsymbol{\theta} \in \mathbb{C}^r$. In order to avoid misinterpretation, we define the new sequence of functions $\{d_n(\boldsymbol{\theta})\}_{n\geqslant 2}$ by setting

$$d_n(\boldsymbol{\theta}) := D_n[w_{\boldsymbol{\theta}}], \quad \boldsymbol{\theta} \in \mathbb{C}^r, \quad n \geqslant 2.$$

From the theory of exponential families and the determinantal identity (6), we obtain that $d_n(\boldsymbol{\theta}) = \tilde{\lambda}_n(\boldsymbol{\theta})(2\pi)^{-n}/n!$ is an entire function for every $n \geqslant 2$. Additionally, the sequence $\{d_n(\boldsymbol{\theta})\}_{n\geqslant 2}$ is locally uniformly bounded as shown next.

Lemma 4.1. *Assume that the conditions of Theorem* 3.1 *are satisfied. Then the sequence* $\{d_n(\boldsymbol{\theta})\}_{n\geqslant 2}$ *of entire functions on* $\mathbb{C}^r$ *is locally uniformly bounded, i.e. for every compact set* $K \subset \mathbb{C}^r$ *there exists a constant* $C_K > 0$ *such that*

$$|d_n(\boldsymbol{\theta})| \leqslant C_K, \quad \forall \boldsymbol{\theta} \in K.$$

Proof. From the determinantal identity (6) we have

$$d_n(\boldsymbol{\theta}) = \frac{1}{(2\pi)^n n!} \int\limits_{[0,2\pi)^n} |\Delta(\boldsymbol{\zeta})|^2 \prod_{k=1}^{n} w_{\boldsymbol{\theta}}(e^{i\zeta_k})\, d\boldsymbol{\zeta}, \quad \boldsymbol{\theta} \in \mathbb{C}^r, \tag{15}$$

and consequently

$$|\, d_n(\boldsymbol{\theta})| \leqslant \frac{1}{(2\pi)^n n!} \int\limits_{[0,2\pi)^n} |\Delta(\boldsymbol{\zeta})|^2 \prod_{k=1}^{n} |w_{\boldsymbol{\theta}}(e^{i\zeta_k})|\, d\boldsymbol{\zeta}, \quad \boldsymbol{\theta} \in \mathbb{C}^r.$$

Since the functions V_j, $1 \leqslant j \leqslant r$, are real-valued, we observe that

$$|w_{\boldsymbol{\theta}}(e^{i\zeta_k})| = \exp\Big\{ \sum_{j=1}^{r} \mathrm{Re}(\theta_j) V_j(e^{i\zeta_k}) \Big\}.$$

From the last equality it follows that it is enough to prove the local uniform bound for the sequence $\{d_n(\boldsymbol{\theta})\}_{n\geqslant 2}$ with respect to the real-valued symbols $w_{\boldsymbol{\theta}}$ with parameter $\boldsymbol{\theta} \in \mathbb{R}^r$. As noticed in [25], for the real-valued symbol $w_{\boldsymbol{\theta}}$, we have from Szegő's theorem that

$$\frac{D_n[w_{\boldsymbol{\theta}}]}{D_{n-1}[w_{\boldsymbol{\theta}}]} = \min_{p \in \mathscr{P}_{n-1}} \int\limits_0^{2\pi} |p(e^{i\zeta})|^2\, w_{\boldsymbol{\theta}}(e^{i\zeta}) d\zeta, \qquad n \geqslant 2, \tag{16}$$

where $\mathscr{P}_n$ is the set of all polynomials of degree not exceeding n and with leading coefficient equal to 1. Observe that the left-hand side in (16) is non-increasing in n and it tends to 1 as $n \to \infty$ because $\lim\limits_{n\to\infty} D_n[w_{\boldsymbol{\theta}}]$ exists. Thus, for every $\boldsymbol{\theta}$, the sequence $\{D_n[w_{\boldsymbol{\theta}}]\}_{n\in\mathbb{N}}$ increases to its limit

$$\exp\Big\{ \sum_{k\in\mathbb{N}} k \Big(\sum_{j=1}^{r} \theta_j\, \widehat{V}_j(k) \Big) \Big(\sum_{l=1}^{r} \theta_l\, \widehat{V}_l(-k) \Big) \Big\}$$

as n increases. As an immediate consequence of this fact, we obtain that

$$|d_n(\boldsymbol{\theta})| \leqslant \exp\Big\{ \frac{1}{2} \sum_{j=1}^{r} \sum_{l=1}^{r} \theta_j \theta_l\, \langle V_j, V_l \rangle_{1/2} \Big\}, \quad \forall n \in \mathbb{N},\ \boldsymbol{\theta} \in \mathbb{R}^r,$$

where $\langle V_j, V_l \rangle_{1/2}$ is the non-negative bilinear form defined in (11). The function on the right-hand side is bounded on compact sets in $\mathbb{R}^r$, and the lemma is proved. $\square$

This lemma proves part of the next lemma.

Lemma 4.2. *Suppose the conditions of Theorem* 3.1 *are satisfied. Then the sequence* $\{d_n(\boldsymbol{\theta})\}_{n\geqslant 2}$ *of Toeplitz determinants* (15) *with respect to the* (*complex-valued*) *generating function* $w_{\boldsymbol{\theta}}$, *converges locally uniformly in* $\mathbb{C}^r$ *to its limit*

$$d(\boldsymbol{\theta}) = \exp\Big\{\frac{1}{2}\sum_{j=1}^{r}\sum_{l=1}^{r}\theta_j\theta_l\,\langle V_j, V_l\rangle_{1/2}\Big\}. \tag{17}$$

Moreover, all partial derivatives of $d_n(\boldsymbol{\theta})$ *converge locally uniformly to the corresponding derivatives of the limiting function* (17).

Proof. The strong Szegő theorem applied to the symbol $w_{\boldsymbol{\theta}} \in L^1(\mathbb{T})$ implies that

$$\lim_{n\to\infty} d_n(\boldsymbol{\theta}) = \exp\Big\{\frac{1}{2}\sum_{j=1}^{r}\sum_{l=1}^{r}\theta_j\theta_l\langle V_j, V_l\rangle_{1/2}\Big\}, \tag{18}$$

under the conditions of Theorem 3.1.

Since the sequence of analytic functions $\{d_n(\boldsymbol{\theta})\}_{n\geqslant 2}$ is locally bounded (Lemma 4.1) and it converges to an entire function on $\mathbb{C}^r$, it follows from Vitali's theorem for several complex variables that it also converges locally uniformly on $\mathbb{C}^r$. If we apply Weierstrass' theorem for several complex variables to the sequence of Toeplitz determinants $\{d_n(\boldsymbol{\theta})\}_{n\geqslant 2}$, we obtain for an arbitrary multi-index $\boldsymbol{l} = (l_1, \ldots, l_r)$, with $\sum_{j=1}^{r} l_j = l,\ l_j \geqslant 0,\ 1 \leqslant j \leqslant r$, that the derivatives

$$\frac{\partial^l}{\partial\theta_1^{l_1}\ldots\partial\theta_r^{l_r}}\, d_n(\boldsymbol{\theta})$$

exist and converge uniformly on compact sets in $\mathbb{C}^r$ to the corresponding derivatives of the limiting function (17). □

In order to obtain the probabilistic interpretation of the uniform convergence of derivatives in Lemma 4.2, we will consider the cumulant generating function $\psi_n(\boldsymbol{\theta}) = \log\lambda_n(\boldsymbol{\theta})$ corresponding to the density $p_{\boldsymbol{\theta};\,n}(\boldsymbol{\zeta})$ and formulate the following corollary.

Corollary 4.3. *Let* $\{p_{\boldsymbol{\theta};\,n}(\boldsymbol{\zeta})\}_{n\geqslant 2}$ *be the sequence of probability density functions defined in* (3) *and suppose that the conditions of Theorem* 3.1 *are satisfied. Then the sequence of gradients* $\{\nabla\psi_n(\boldsymbol{\theta})\}_{n\geqslant 2}$ *of its cumulant* *generating functions ψ_n converges locally uniformly in $\boldsymbol{\theta} \in \mathbb{R}^r$ to the vector*

$$\Big(\sum_{j=1}^{r} \theta_j \langle V_1, V_j\rangle_{1/2}, \ldots, \sum_{j=1}^{r} \theta_j \langle V_r, V_j\rangle_{1/2}\Big).$$

Proof. The statement follows from Lemma 4.2 applied to compact sets $K \subset \mathbb{R}^r$ and the functions

$$\frac{\partial}{\partial\theta_j}\psi_n(\boldsymbol{\theta}) = \frac{\partial}{\partial\theta_j} d_n(\boldsymbol{\theta})/d_n(\boldsymbol{\theta}), \quad j = 1, \ldots, r. \qquad \square$$

As discussed in the beginning of the section, our aim is to obtain the asymptotic maximum likelihood estimator for the vector parameter $\boldsymbol{\theta} \in \mathbb{R}^r$. Since all preliminary results are stated, we may summarize the discussion in the following theorem. We denote the true parameter value by $\boldsymbol{\theta}_0$.

Theorem 4.4. *Let $\{p_{\boldsymbol{\theta};n}(\boldsymbol{\zeta})\}_{n\geqslant 2}$ be the sequence of probability density functions given in* (3)*. Assume that the conditions of Theorem* 3.1 *are satisfied and that the functions $V_j, j = 1, \ldots, n$, are linearly independent. Then the asymptotic maximum likelihood estimator $\widehat{\boldsymbol{\theta}}^{(n)}$ of $\boldsymbol{\theta}_0$, defined as the unique solution of the system of linear equations*

$$\sum_{l=1}^{r} \widehat{\theta}_l^{(n)} \langle V_j, V_l\rangle_{1/2} = \sum_{k=1}^{n} V_j(e^{i\zeta_k}), \qquad 1 \leqslant j \leqslant r, \tag{19}$$

converges weakly, under the distributions with density $p_{\boldsymbol{\theta}_0,n}$, to $\mathscr{N}(\boldsymbol{\theta}_0, \Sigma^{-1})$, where Σ is as specified in (13)*.*

Proof. The maximum likelihood estimator $\tilde{\boldsymbol{\theta}}^{(n)}$ is defined as the solution of the system of equations

$$\frac{\partial}{\partial\theta_j}\psi_n(\tilde{\boldsymbol{\theta}}^{(n)}) = \sum_{l=1}^{n} V_j(e^{i\xi_l}), \qquad 1 \leqslant j \leqslant r. \tag{20}$$

The convergence in Corollary 4.3 implies that the left-hand side in (20) can be approximated by the expression

$$\sum_{k=1}^{r} \tilde{\theta}_k^{(n)} \langle V_j, V_k\rangle_{1/2}.$$

This justifies the definition of the asymptotic maximum likelihood estimator $\widehat{\boldsymbol{\theta}}^{(n)}$ by the system of linear equations (19).

It also follows that

$$\widehat{\boldsymbol{\theta}}^{(n)} = \boldsymbol{V}^{(n)}(\boldsymbol{\xi})\, \Sigma^{-1}$$

and, consequently, Theorem 3.1 yields the convergence of

$$\widehat{\boldsymbol{\theta}}^{(n)} \xrightarrow{\mathscr{D}} \mathscr{N}(\boldsymbol{\theta}_0, \Sigma^{-1}), \qquad n \to \infty. \qquad \square$$

Remark 4.5. We notice that although $\widehat{\boldsymbol{\theta}}^{(n)}$ is asymptotically unbiased and follows asymptotically a normal distribution, it is not a consistent estimator, as mentioned before. This fact can be explained by the behavior of the Fisher information matrix $I_n(\boldsymbol{\theta})$ corresponding to the density $p_{\boldsymbol{\theta};\, n}(\boldsymbol{\zeta})$. Namely,

$$I_n(\boldsymbol{\theta}) = \left\| \frac{\partial^2}{\partial\theta_j \partial\theta_l} \psi_n(\boldsymbol{\theta}) \right\|_{j,l=1,\dots,r}$$

converges to the constant matrix Σ defined in (13), and therefore, the information contained in the observed sample $(\zeta_1, \dots, \zeta_n)$ of exchangeable random variables with probability density $p_{\boldsymbol{\theta};\, n}(\boldsymbol{\zeta})$ remains bounded as n increases. This fact can be given an interpretation in the framework of the theory of log-potential gases. Our procedure involves estimation of parameters of an external field for a Coulomb gas with Boltzmann factor (3). While the number of particles in the Coulomb gas increases, the potential of an external field remains unchanged. As a consequence, the information about the potential contained in the observation $\boldsymbol{\zeta}$ is bounded. We conclude that consistent maximum likelihood estimation of the parameter $\boldsymbol{\theta}$ is not possible, unless the dimension parameter n is introduced into the weight function $w_{\boldsymbol{\theta}}$ defined in (4). In such a setting the potential of an external field is assumed to be proportional to the number of particles in the Coulomb gas system. Below, we consider the model (2) with varying weight and the third-order phase transition, and show that under such a model, the asymptotic maximum likelihood estimator of a concentration parameter γ has the variance of order $O(1/n^2)$.

§5. The third-order phase transition model

In this section we consider the model from two-dimensional lattice gauge theories which was analyzed heuristically by the steepest descent method in Gross, Witten [19]. The same model arose in the studies of the length of the longest increasing subsequence in a random permutation. Its properties were rigorously analyzed in a series of papers including Johansson [26], Baik, Deift, Johansson [3], Baik, Deift, Rains [4] and Widom [44]. It was shown that the ensemble exhibits the third-order phase transition at $\gamma = 1$.

Formally, the model is defined by the joint probability density of n eigenphases equal to

$$p_{\gamma;\,n}(\boldsymbol{\zeta}) = C_n^{-1}(\gamma)|\Delta(\boldsymbol{\zeta})|^2 \exp\Big\{\gamma n \sum_{i=1}^{n} \cos\zeta_i\Big\}, \qquad \boldsymbol{\zeta} \in [0, 2\pi]^n, \tag{21}$$

where $w_\gamma(e^{i\zeta}) := \exp(\gamma\cos\zeta)$ and where (abusing the previous notation a bit) $C_n(\gamma) = (2\pi)^n n!\, D_n[w_{\gamma n}]$ is the normalizing constant and $\gamma > 0$.

The following lemma, due to Gross, Witten [19] and Johansson [26, Lemma 2.1], describes the asymptotic behavior of the free energy $f_n(\gamma) = n^{-2}\log C_n(\gamma)$.

Lemma 5.1 (Gross, Witten [19]). *If $f_n(\gamma) = n^{-2}\log C_n(\gamma)$, then*

$$\lim_{n\to\infty} f_n(\gamma) = f(\gamma) = \begin{cases} \frac{\gamma^2}{4}, & \text{if } \ 0 \leqslant \gamma \leqslant 1, \\ \gamma - \frac{1}{2}\log\gamma - \frac{3}{4}, & \text{if } \ 1 < \gamma. \end{cases}$$

Remark 5.2. The limiting function $f(\gamma)$ in Lemma 5.1 is not analytic. Its derivative $d^3f/d\gamma^3$ is discontinuous at $\gamma = 1$, thus a third-order phase transition occurs at this point. The asymptotic eigenvalue distribution is supported on the whole unit circle for $\gamma < 1$, whereas for $\gamma > 1$ its support is a subset of $\mathbb{T}$, details are contained in Johansson [26].

The following lemma shows that the limit in Lemma 5.1 can be used to prove the asymptotic normality of the sufficient statistics $T_n(\boldsymbol{Z})$ defined in (22).

Lemma 5.3. *Let $0 < \gamma < 1$ and $\boldsymbol{Z} = (Z_1, \ldots, Z_n)$ be a sample with the joint probability density function* (21)*. Then the centered sufficient statistic*

$$T_n(\boldsymbol{Z}) = \sum_{k=1}^{n} \cos Z_k - n\gamma/2 \tag{22}$$

converges in distribution to the normal distribution with mean 0 *and variance* 1/2*.*

Proof. The moment generating function of the statistics $T_n(\boldsymbol{Z})$ is

$$\begin{aligned} M_n(s) &= C_n^{-1}(\gamma)e^{-s\gamma n/2} \int\limits_{[0,2\pi]^n} |\Delta(\boldsymbol{\zeta})|^2 \exp\Big\{(\gamma n + s)\sum_{k=1}^{n}\cos\zeta_k\Big\} d\boldsymbol{\zeta} \\ &= e^{-s\gamma n/2}\frac{D_n[w_{\gamma n+s}]}{D_n[w_{\gamma n}]} = e^{-s\gamma n/2}\frac{C(\gamma + \frac{s}{n})}{C(\gamma)} \end{aligned}$$

for $s \in \mathbb{R}$. Let $0 < \epsilon < 1 - \gamma$ and consider the sequence of functions $\{M_n(s)\}_{n \in \mathbb{N}}$ where s belongs to the interval

$$s \in I_\gamma = [-\min(\gamma, 1-\gamma), \min(\gamma, 1-\gamma)].$$

From the definition of I_γ it follows that $0 \leqslant \gamma + s/n \leqslant 1$ for every $s \in I_\gamma$, $n \geqslant 1$. It has been shown in Johansson [26, Lemma 3.3], that for every $\eta > 0$ there exists a constant $K(\eta) > 0$ such that

$$\Big| \log C_n(\delta) - \frac{n^2\delta^2}{4} \Big| \leqslant \frac{K(\eta)}{n},$$

for arbitrary $n \geqslant 1, \delta \in [0, 1-\eta]$. Consequently, with $\delta \in \{\gamma, \gamma + \frac{s}{n}\}$ and $\eta = \epsilon$

$$e^{-K(\epsilon)/n} \leqslant C_n(\gamma)\, e^{-\gamma^2 n^2/4} \leqslant e^{K(\epsilon)/n}, \quad n \geqslant 1,$$

and

$$e^{-K(\epsilon)/n} \leqslant C_n(\gamma + s/n)\, e^{-(\gamma + \frac{s}{n})^2 n^2/4} \leqslant e^{K(\epsilon)/n}, \qquad n \geqslant 1,$$

uniformly for $\gamma \in [0, 1-\epsilon)$. Thus, we have the following expansion

$$\begin{aligned} M_n(s) &= \exp\Big\{ -\frac{sn\gamma}{2}\ Big\} \exp\Big\{ \frac{n^2}{4}\Big(\gamma + \frac{s}{n}\Big)^2 - \frac{n^2}{4}\gamma^2 + O\Big(\frac{1}{n}\Big)\Big\} \\ &= \exp\Big\{ \frac{s^2}{4} + O\Big(\frac{1}{n}\Big)\Big\}, \qquad s \in I_\gamma. \end{aligned}$$

Therefore,

$$\lim_{n\to\infty} M_n(s) = \exp(s^2/4), \quad s \in I_\gamma,$$

where the function on the right-hand side is the moment generating function of the normal distribution with mean 0 and variance 1/2. The asymptotic normality of the statistics $T_n(\boldsymbol{Z})$ follows. $\square$

Corollary 5.4. *If the conditions of Lemma* 5.3 *are satisfied, the asymptotic maximum likelihood estimator of the parameter* γ_0

$$\widehat{\gamma}_n = \frac{2}{n}\sum_{i=k}^{n} \cos Z_k$$

is asymptotically unbiased and

$$n(\widehat{\gamma}_n - \gamma_0) \xrightarrow{\mathscr{D}} \mathcal{N}(0, 2)$$

as $n \to \infty$.

Proof. The maximum likelihood estimator $\tilde{\gamma}_n$ of the parameter γ is obtained from the equation

$$\frac{d}{d\gamma}\log C_n(\tilde{\gamma}_n) = n\sum_{k=1}^{n}\cos\zeta_k. \tag{23}$$

Since it was proved in Johansson [26, Lemma 3.3], that for $0 < \epsilon < 1$ there exists a constant $K > 0$, such that

$$\left|\frac{d}{d\gamma}\log C_n(\gamma) - \frac{n^2\gamma}{2}\right| \leqslant \frac{K}{n},$$

for every $n \geqslant 1, \gamma \in (0, 1-\epsilon]$, we replace the equation (23) by

$$\frac{n^2\widehat{\gamma}_n}{2} = n\sum_{k=1}^{n}\cos\zeta_k,$$

which defines the asymptotic maximum likelihood estimator as before. This leads to the estimator

$$\widehat{\gamma}_n = \frac{2\sum\limits_{k=1}^{n}\cos\zeta_k}{n}.$$

We notice that

$$\mathrm{E}_n[\widehat{\gamma}_n - \gamma_0] = \mathrm{E}_n[2T_n(\boldsymbol{Z})/n] \to 0, \quad n \to \infty,$$

and

$$n(\widehat{\gamma}_n - \gamma_0) \xrightarrow{\mathscr{D}} \mathcal{N}(0,2), \qquad n \to \infty,$$

proving the corollary. □

§6. Open questions

Remark 6.1. The case $\gamma > 1$ is not considered in Section 5. In this case the maximum likelihood equation (23) should be replaced by the asymptotic expression

$$1 - \frac{1}{2\tilde{\gamma}_n} = \frac{\sum\limits_{i=1}^{n}\cos\zeta_i}{n}$$

which follows from Lemma 5.1 and leads us to the equation which defines the asymptotic maximum likelihood estimator $\widehat{\gamma}_n$. Asymptotic properties of $\widehat{\gamma}_n$ can be derived using the same methods as here. Intuitively, the asymptotic distribution of the sufficient statistics

$$T_n(\boldsymbol{Z}) = \sum_{k=1}^{n} \cos Z_k - n\Big(1 - \frac{1}{2\gamma}\Big)$$

is a consequence of the following convergence of the moment generating functions

$$\lim_{n\to\infty} M_n(s) = \exp\Big[\frac{s^2}{2}\,\frac{1}{2\gamma^2}\Big], \quad s \in [1-\gamma, \gamma-1], \tag{24}$$

and consequently, the asymptotic distribution of $T_n(\boldsymbol{Z})$ is normal with expectation 0 and variance $(2\gamma^2)^{-1}$. However, to prove the convergence (24) rigorously, one needs to establish that for $\gamma > 1$, the rate of convergence to the limiting function in Lemma 5.1 is of order $o(n^{-2})$.

Remark 6.2. Various extensions of questions addressed in this paper are possible. The Bessel weight in (21) could be replaced by any other varying weight on the unit circle. Phase transition phenomena similar to those appearing in the case of Hermitian matrices (see [13]) are expected to occur. We should note here that while models with varying exponential weights for Hermitian matrices are extensively studied (see [8, 13] and references therein), respective circular models are only starting to be considered (see [30]). We believe that these models may become an extremely rich research area. The models show high mathematical complexity, many different analytic methods are involved, and the models have a wide range of applications in different branches of mathematics and physics. The question of non-parametric estimation in the framework of generalized circular unitary ensembles with varying weights remains open. Finally, similar generalizations of circular orthogonal and symplectic ensembles are to be considered.

References

1. M. Adler, P. van Moerbeke, *Recursion relations for unitary integrals, combinatorics and the Toeplitz lattice.* — Comm. Math. Phys. **237**, No. 3 (2003), 397–440.
2. T. W. Anderson, *An Introduction to Multivariate Statistical Analysis.* Wiley Series in Probability and Statistics. Wiley-Interscience [John Wiley & Sons], Hoboken, NJ, third edition, 2003.
3. J. Baik, P. Deift, K. Johansson, *On the distribution of the length of the longest increasing subsequence of random permutations.* — J. Amer. Math. Soc. **12**, No. 4 (1999), 1119–1178.

4. J. Baik, P. Deift, E. Rains, *A Fredholm determinant identity and the convergence of moments for random Young tableaux.* — Comm. Math. Phys. **223**, No. 3 (2001), 627–672.
5. J. Baik, *Circular unitary ensemble with highly oscillatory potential.* ArXiv:1306.0216v2, June 13, 2013.
6. Z. Bai, J. Silverstein, *Spectral Analysis of Large Dimensional Random Matrices.* Second edition. Springer Series in Statistics. Springer, New York, 2010.
7. C. W. J. Beenakker, *Universality in the random-matrix theory of quantum transport.* — Phys. Rev. Lett. **70** (1993), 1155–1158.
8. P. M. Bleher, A. R. Its, *Asymptotics of the partition function of a random matrix model.* — Ann. Inst. Fourier (Grenoble) **55**, No. 6 (2005), 1943–2000.
9. A. Borodin, A. Okounkov, *A Fredholm determinant formula for Toeplitz determinants.* — Integral Equations Operator Theory **37**, No. 4 (2000), 386–396.
10. A. Borodin, G. Olshanski, *Harmonic analysis on the infinite-dimensional unitary group and determinantal point processes.* — Ann. Math. (2) **161**, No. 3 (2005), 1319–1422.
11. A. Borodin, *Discrete gap probabilities and discrete Painlevé equations.* — Duke Math. J. **117**, No. 3 (2003), 489–542.
12. L. D. Brown, *Fundamentals of Statistical Exponential Families with Applications in Statistical Decision Theory.* Institute of Mathematical Statistics Lecture Notes Monograph Series 9. Institute of Mathematical Statistics, Hayward, CA, 1986.
13. P. A. Deift, *Orthogonal Polynomials and Random Matrices: a Riemann–Hilbert Approach.* Courant Lecture Notes in Mathematics 3. New York University Courant Institute of Mathematical Sciences, New York, Amer. Math. Soc, Providence RI, 1999.
14. P. Diaconis, S. N. Evans, *Linear functionals of eigenvalues of random matrices.* — Trans. Amer. Math. Soc. **353**, No. 7 (2001), 2615–2633.
15. P. Di Francesco, P. Ginsparg, J. Zinn-Justin, *2D gravity and random matrices.* — Phys. Rep. **254**, No. 1–2 (1995), 133.
16. I. Dumitriu, A. Edelman, *Global spectrum fluctuations for the β-Hemite and β-Lagurre ensembles via matrix models.* — J. Math. Phys. **47**, No. 6:063302, (2006).
17. F. J. Dyson, *Statistical theory of the energy levels of complex systems.* — I. J. Math. Phys. **3** (1962), 140–156.
18. P. Forrester, *Log-Gases and Random Matrices.* London Math. Soc. Monographs Series, **34**. Princeton University Press, Princeton, NJ, 2010.
19. D. J. Gross, E. Witten, *Possible third-order phase transition in the large-n lattice gauge theory.* — Phys. Rev. D **21**, No. 2 (1980), 446–453.
20. F. Haake, *Quantum Signatures of Chaos.* Springer Series in Synergetics. Springer-Verlag, Berlin, second edition, 2001. With a foreword by H. Haken.
21. I. A. Ibragimov, *A theorem of Gabor Szegő.* — Mat. Zametki **3** (1968), 693–702.
22. M. E. H. Ismail, N. S. Witte, *Discriminants and functional equations for polynomials orthogonal on the unit circle.* — J. Approx. Theory **110**, No. 2 (2001), 200–228.
23. R. A. Jalabert, J. Pichard, *Quantum mesoscopic scattering: Disordered systems and Dyson circular ensembles.* — J. Phys. I France **5** (1995), 287–324.

24. S. R. Jammalamadaka, A. SenGupta, *Topics in Circular Statistics.* Series on Multivariate Analysis 5. World Scientific Publishing Co. Inc., River Edge, NJ, 2001.
25. K. Johansson, *On random matrices from the compact classical groups.* — Ann. Math. (2) **145**, No. 3 (1997), 519–545.
26. K. Johansson, *The longest increasing subsequence in a random permutation and a unitary random matrix model.* — Math. Res. Lett. **5**, No. 1–2 (1998), 63–82.
27. K. Johansson, *On fluctuations of eigenvalues of random Hermitian matrices.* — Duke Math. J. **91**, No. 1 (1998), 151–204.
28. J. V. José, R. Cordery, *Study of a quantum fermi-acceleration model.* — Phys. Rev. Lett. **56**, No. 4 (1986), 290–293.
29. R. Matic, *Estimation problems related to random matrix exsembles.* PhD dissertation, Georg-August-Universität Göttingen, 2006.
30. K. T.-R. McLaughlin, P. D. Miller, *The $\overline{\partial}$ steepest descent method and the asymptotic behavior of polynomials orthogonal on the unit circle with fixed and exponentially varying non-analytic weights.* IMRP Int. Math. res. Pap., pages Art. ID 48674, 1–77, 2006.
31. M. L. Mehta, *Random Matrices.* Pure and Applied Mathematics 142 (Amsterdam). Elsevier/Academic Press, Amsterdam, third edition, 2004.
32. R. J. Muirhead, *Aspects of Multivariate Statistical Theory.* Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, 1982.
33. K. E. Muttalib, M. E. H. Ismail, *Impact of localization on Dysons circular ensemble.* — J. Phys. A: Math. Gen. **28** (1995), L541–L548.
34. T. Nagao, M. Wadati, *An integration method on generalized circular ensembles.* — J. Phys. Soc. Japan **61**, No. 6 (1992), 1903–1909.
35. V. Periwal, D. Shevitz, *Unitary-matrix models as exactly solvable string theories.* — Phys. Rev. Lett. **64**, No. 12 (1990), 1326–1329.
36. P. Rambour, A. Seghier, *The generalised Dyson circular unitary ensemble: asymptotic distribution of the eigenvalues at the origin of the spectrum.* — Integral Equations Operator Theory **69**, No. 4 (2011), 535–555.
37. B. Simon, *Orthogonal Polynomials on the Unit Circle.* Part 1. Classical Theory. American Mathematical Society Colloquium Publications 54. American Mathematical Society, Providence, RI, 2005.
38. A. Soshnikov, *The central limit theorem for local linear statistics in classical compact groups and related combinatorial identities.* — Ann. Probab. **28**, No. 3 (2000), 1353–1370.
39. G. Szegő, *Orthogonal Polynomials.* American Mathematical Society, Providence, R.I., fourth edition, 1975. American Mathematical Society, Colloquium Publications, Vol. XXIII.
40. C. A. Tracy, H. Widom, *Random unitary matrices, permutations and Painlevé.* — Comm. Math. Phys. **207**, No. 3 (1999), 665–685.
41. A. M. Tulino, S. Verdù, *Random matrix theory and wireless communications.* — Foundation and Trends in Communications and Information Theory **1**, No. 1 (2004), 1–182.
42. H. Weyl, *Theorie der Darstellungen kontinuierlicher halbeinfacher Gruppen durch lineare Transformationen.* I. — Math. Z. **23**, No. 1 (1925), 271–309.

43. H. Weyl, *Theorie der Darstellungen kontinuierlicher halbeinfacher Gruppen durch lineare Transformationen.* II. — Math. Z. **24**, No. 1 (1926), 328–376.
44. H. Widom, *On convergence of moments for random Young tableaux and a random growth model.* — Int. Math. Res. Not. **9** (2002), 455–464.
45. E. P. Wigner, *On the statistical distribution of the widths and spacings of nuclear resonance levels.* — Proc. Cambridge. Phil. Soc. **47** (1951), 790–798.
46. E. P. Wigner, *Characteristic vectors of bordered matrices with infinite dimensions.* — Ann. Math. (2) **62** (1955), 548–564.
47. E. P. Wigner, *On the distribution of the roots of certain symmetric matrices.* — Ann. Math. (2) **67** (1958), 325–327.
48. J. Wishart, *The generalized product moment distribution in samples from a normal multivariate population.* — Biometrika **20** (1928), A:32–43.

Georg-August-Universität Göttingen

Поступило 30 сентября 2015 г.

The Pennsylvania State University
E-mail: mhd13@psu.edu

Steklov Institute of Mathematics, St. Petersburg